INTEGRITY MONITORING IN SNAPSHOT AND RECURSIVE ESTIMATION ALGORITHMS FOR MARITIME APPLICATIONS

Luís Lança, Michailas Romanovas, Ralf Ziebold

German Aerospace Centre (DLR)
Institute of Communications and Navigation
Kalkhorstweg 53, 17235, Neustrelitz, Germany
Email:{Luis.Lanca, Michailas.Romanovas, Ralf.Ziebold}@dlr.de

1 INTRODUCTION

With the advent of Global Navigation Satellite Systems (GNSS) the positioning fixing by electronic means has become a key component for a variety of applications and can be considered as a main source for maritime Positioning, Navigation and Timing (PNT) data provision. Navigation information systems compliant with IMO (International Maritime Organization) regulations, like the Automatic Identification System (AIS) integrate the GNSS based positioning information [1]. The complexity of the navigation process and the associated modern technologies involved had led the IMO to start the so called e-Navigation Initiative which aims the harmonization of systems and standards [2]. In order to support the strategic goal of resilient PNT data provision, the German Aerospace Center has developed a PNT Unit concept, where all navigational sensors are used in a combined way. To confirm the performance under real operational conditions an operational prototype of such a PNT unit was realized. Here the integrity monitoring (IM) algorithms are responsible for the evaluation of the events or conditions that have the potential to cause or to contribute to hazardously misleading information (HMI).

The presented work addresses the performance of the IM in a snapshot and a Recursive Bayesian Estimation (RBE) algorithms within maritime application scenarios, where IM form a mechanism to protect in a timely fashion from large position errors in the presence of failures or non-scheduled events. A key component of the IM is the fault detection and exclusion (FDE) function which is responsible for the elimination of faulty data sources or measurements (so-called outliers).

The report concentrates on analysis of the least-squares (LS) residual FDE algorithm and its impact on the navigation quality. The snapshot LS-based FDE is also tested against an Extended Kalman Filter (EKF) FDE approach both for non-weighted and weighted range observation models. The real sensor data collected during a typical port operation scenario is tested for the outlier rejection performance by adding artificial errors of varying amplitude and evaluating the performance of the developed FDE schemes. In order to make the analysis statistically relevant, a Monte Carlo simulation has been performed, where the true GNSS measurements with only few inherent outliers with real noise and signal impairments (e.g. multipath effects) and the characteristics from the real environment are contaminated with the artificial outliers.

The Snapshot Least-Squares Residuals (LSR) Receiver Autonomous Integrity Monitoring (RAIM) developed by the aviation community [7] or the statistical reliability testing techniques developed by the geodetic community [10] are classic references for LSR FDE algorithms. The approaches make use of the measurements redundancy to check the relative consistency among the residuals in order to detect the most likely measurement fault. Most of the approaches are based on the test against assumed statistics. Although the classical methods mainly use snapshot techniques, some works [8] have been reported on introducing the FDE algorithms for RBE techniques [13].

The CNo, which is the ratio of the received carrier (i.e. the signal) power to noise density can be employed as signal quality indicator for a more precise positioning [14]. Although the CNo adaptive weighting models perform as good as elevation angle-based methods for clear sky signal environment, they often outperform the latter in signal degraded environments [5].

2 METHODS

The navigation applications are usually formulated as state estimation problems using a combination of the measurements from multiple sensors with complementary properties. A desired set of the parameters to be estimated from the noisy measurements usually includes the object's position, velocity, etc. Here we utilize a well-established estimation strategy based on the RBE framework, while a classical LS solution can be considered as a non-recursive memory-less approach. The classical RBE cycle is performed in two steps:
Prediction The *a priori* probability is calculated from the last *a posteriori* probability using the process model.
Correction The *a posteriori* probability is calculated from the *a priori* probability using the measurement model and the current measurement.
In practice, however, the methods formulated with probability densities do not scale up very well and can quickly become intractable even for the estimation problems of reasonable dimensionality. Various implementations of RBE algorithms differ in the way the probabilities are represented and transformed in the process and measurement models [11, 4]. If the models are linear and the probabilities are Gaussian, the linear KF is an efficient and optimal solution of the estimation problem. Unfortunately, most of the real-world systems are rather nonlinear and modifications to the linear KF have been developed to deal with nonlinear models.

The EKF is one of the most popular nonlinear estimators and is historically considered as a standard within the engineering community. In EKF the nonlinear models are linearized about the current estimate using the Taylor series expansion, where the state transition A and observation matrices H are replaced by the corresponding Jacobians. The system at every time t_k is represented by the state x_k and an associated covariance P_k with the rest of the filtering scheme being essentially identical to that of the classical linear KF. Although the EKF inherits many advantages of the KF such as limited computational costs and clear filtering structure, it still suffers from two main problems. Firstly, the performance of the estimator strongly depends on the validity of the linearized model assumption and can become inaccurate and lead to filter instabilities if these assumptions are violated. Secondly, the required Jacobians can be potentially difficult or even impossible to derive if dynamical models involves complex approximation coefficients and/or discontinuities.

For the LS estimation we follow a classical approach with linearization of the measurement function at each epoch t_k around a point x_0 and finding the correction factor $\delta\hat{x}$ using [3]:

$$\delta\hat{x} = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} \delta z, \tag{1}$$

and the iterative update of the initial estimates $x_i = x_{i-1} + \delta\hat{x}_i$, where δz is the misclosure vector. The usual LS GNSS solution involves four unknowns and if there are five or more observations available (i.e. $m > n$), the redundancy could be used to check the consistency among the full set of the observed measurements [7]. This forms a fundamental principle for the fault detection using LSR using the measurement residual norm $\|r\|$:

$$\hat{r} = z - H\hat{x} = (I - H(H^T R^{-1} H)^{-1} H^T R^{-1}) z. \tag{2}$$

The test statistics is based on the normalized estimated residual vector norm $\|\hat{r}\|^2 = \hat{r}^T R^{-1} \hat{r}$. In the fault-free case the individual residuals follow $\mathcal{N}(0,1)$, while the norm follows a central Chi-Square distribution with $m - n$ degrees of freedom (DoF). The LSR FDE is based on the hypothesis testing which compares the test statistics with the given threshold. The weighted LSR RAIM statistics is defined as $ts = \sqrt{\|\hat{r}\|^2}$, which for a fault free case is tested with a threshold Th for a given probability of false alarm (Pfa) and DoF [12]. A common procedure consists of fixing Pfa according to the application requirements and letting the threshold vary with the measurement redundancy:

$$\text{Global-Hypothesis-Test} = \begin{cases} H_0 & \text{if } t_s \leq T_h, \\ H_1 & \text{if } t_s > T_h. \end{cases} \tag{3}$$

If the null hypothesis is rejected, an inconsistency in the tested measurements is assumed and the fault source should be identified and excluded. An exclusion method based on a local test is adopted [8, 5] which assesses the standardized residuals defined as follows:

$$w_i = \left| \frac{\hat{r}_i}{\sqrt{(C_{\hat{r}})_{i,i}}} \right|, i = [1, \ldots, n], \tag{4}$$

where $C_{\hat{r}}$ is the covariance matrix for residuals $C_{\hat{r}} = R - H(H^T R^{-1} H)^{-1} H^T$. Assuming one fault at a time, the w_i is tested using the quantile of a normal distribution equal to the Pfa. In the local test the residual is excluded if w_i exceeds the test threshold. Again, the local test assumes the residuals to be $\mathcal{N}(0,1)$ in the fault-free case:

$$\text{Local-Hypothesis-Test} = \begin{cases} H_{0,i} & \text{if } w_i \leq z_{1-Pfa/2}, \\ H_{1,i} & \text{if } w_i > z_{1-Pfa/2}, \end{cases} \tag{5}$$

where z is the quantile of the standard normal distribution. Only the largest w_i is tested against the $H_{0,i}$ and the measurement k is selected as a candidate to be excluded if and only if $w_k \geq w_i,\ \forall i$ and $w_k > z_{1-Pfa/2}$. The KF predicted innovation vector $e_k = z_k - H\hat{x}_k^-$ is considered as an indication of the amount of the information introduced to the system and its norm can be used as a quality indicator. For a fault free situation the norm follows a central Chi Squared distribution with $m - n$ DoF. The global test statistics is given by [8]: $ts_{kf} = \sqrt{\hat{e_k}^T C_{e,k}^{-1} \hat{e_k}}$ with $C_e = H_k P_k^- H_k^T + R_k$. The tests are performed following the same two-step LSR procedure [5]. Firstly, the global test checks the consistency and if some inconsistency is detected, the scheme performs a local test. The latter is recursively applied whenever a fault is detected until no more faults are found [8].

As the basis for the adaptive pseudorange measurement noise covariance model we have adopted the following expression:

$$\sigma^2 = a + b \cdot 10^{-\frac{CNo-c}{10}}, \tag{6}$$

with three approximation parameters a, b and c. In the expressions above the CNo is the measured carrier to noise density (in dB-Hz) for a particular pseudorange observation. Obviously, different fitting functions can be constructed, but one does not expect a fundamentally different behavior of these models as all of the perform similarly by assigning a larger noise covariance for the measurements with lower CNo.

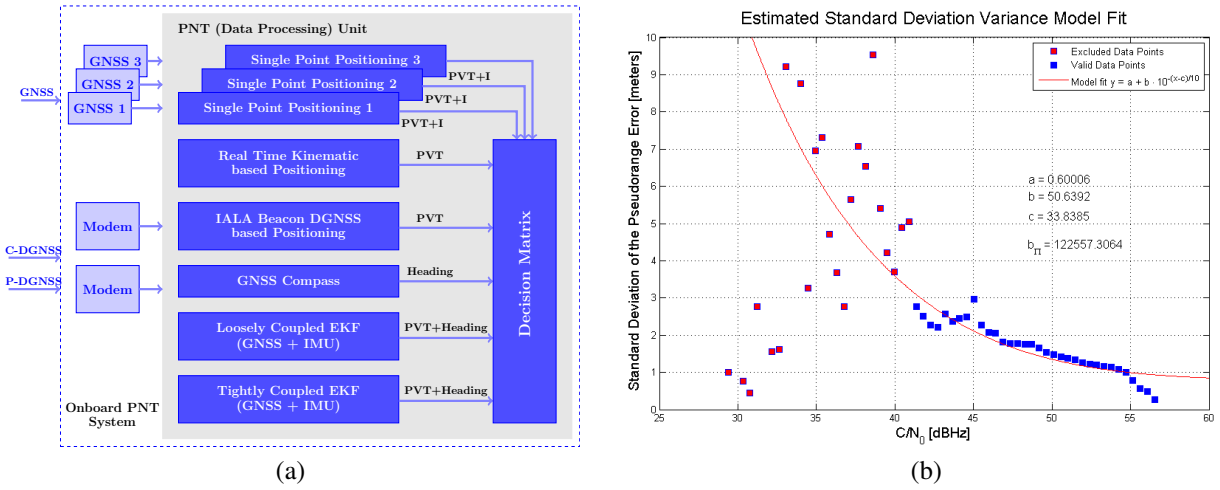
3 SYSTEM

The vessel Baltic Taucher II has been used as a base platform for the PNT unit development and was equipped with three dual frequency GNSS antennas and receivers (Javad Delta), an IMU (iMar IVRU FCAI), a gyrocompass, a Doppler speed log and an echo sounder. The IALA Beacon antenna and receiver were employed for the reception of IALA Beacon DGNSS corrections and a UHF modem for the reception of RTK correction data from our Maritime Ground Based Augmentation System (MGBAS) station located in the port of Rostock [6]. The MGBAS reference station provides GPS code and phase corrections with 2 Hz update rate for both L1 and L2 frequencies. These correction data are used for a high accurate RTK positioning onboard the vessel as well as for generation of the post processed reference trajectory with the Javad Justin program.

All sensor data are provided either directly via Ethernet or via a serial to Ethernet adapter to a Box PC where the measurements are processed in real time and additionally stored in a SQlite3 database along with the corresponding time stamps. This setup enables a record and replay functionality for further processing of the original sensor data in our laboratory. Here Fig. 1(a) provides a sketch of the proposed on-board PNT system, which consists of a modular hardware platform and a Real-Time software Framework (RTF) implemented in ANSI-C++. Here several processing channels can be implemented to increase the robustness of the system using multiple sensors as well as different algorithmic approaches for the PNT data estimation. The data used in this paper was recorded on the 01/09/2014 in a quasi-static scenario where the vessel Baltic Taucher II was moored at his home port 'Alter Fischereihafen' on the river Warnow close to the Rostock port with weak wind and little waves. The evaluation is based on data from the mid ship antenna and some shadowing effects from the main mast and the obstacles in the surrounding can be expected. The chosen environment represents a typical maritime port application.

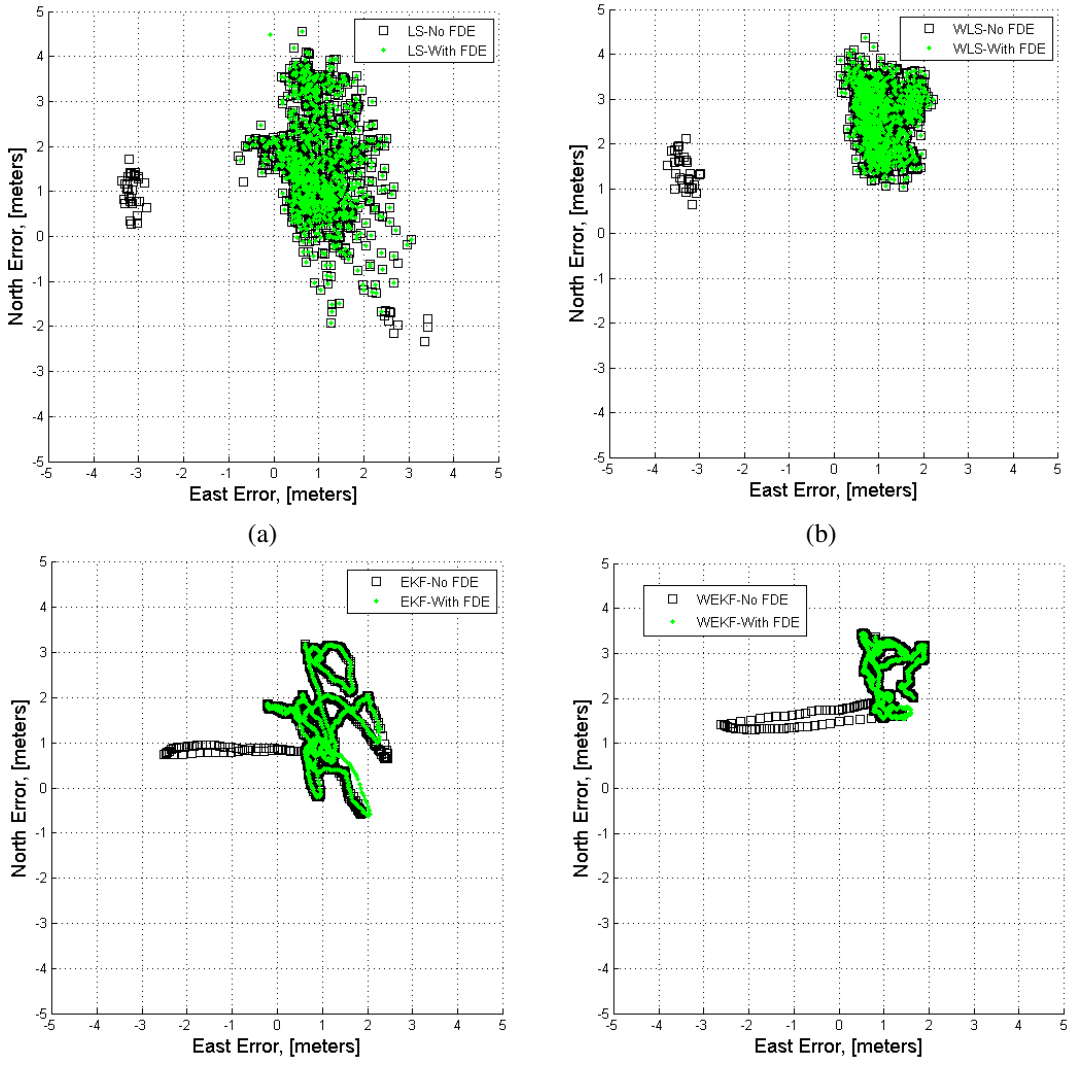
4 RESULTS

In order to evaluate the performance of the methods, a realistic adaptive pseudorange noise model had to be extracted. The experimental data have been obtained from a reference receiver of known position over 24 hours using broadcast ionospheric and tropospheric corrections with the error statistics computed by analyzing the differences between the expected and the observed ranges. The obtained data have been binned according to the associated CNo values and for each



(a) (b)

Figure 1: The concept of the developed on-board PNT unit (a) and the experimental data for PR noise model and model fit results (b).



(a) (b)

(c) (d)

Figure 2: Performance of the FDE in terms of horizontal position accuracy for the non-weighted LS approach (a), weighted LS approach (b), non-weighted EKF (c) and weighted WEKF (d) when an outlier of 15 meters is applied.

bin a variance could be estimated. Note that in this simplified approach only variance was modeled as a function of the signal quality and the non-zero mean offset was ignored. In practice, however, the latter could be also modeled, e.g, by increasing even further the associated variance. Figure 1(b) shows the experimental results and the extracted model using a nonlinear least squares fit. The points with lower CNo values have been manually excluded and the effective fit was found to the values larger than 40 dB-Hz. The residuals with CNo lower than 40 dB-Hz have often insufficient statistics and therefore cannot serve for the model extraction. Moreover, the performance of the DLL correlator in the GNSS receiver is often poor for low CNo values and the obtained values are simply not representative. The extracted model was used in weighted methods as described below and allows to construct both adaptive LS and KF schemes. Similar model can be also extracted for the Doppler measurements in order to use the rate observations even from the satellites with low CNo values. Note that the experimental data show also a small noise for CNo larger than 55 dB-Hz. The observed values are close or even smaller than the correlator base noise level and are probably caused by the insufficient sample size for larger CNo values. Note that this effect has been effectively eliminated from the fit model as the parameter a is almost 60 cm which is close to the rough theoretical calculations for the associated hardware. This also allows us not to exclude these points manually as both LS and KF algorithms have shown relatively low sensitivity to a small variations in variance models.

Figure 2 shows the results of the implemented FDE on the performance of a classical LS solution (a), the LS solution with adaptive weighting model (b), direct EKF with equally weighted measurements ($\sigma = 2$) meters (c) and adaptively weighted DEKF (d) in the presence of an artificially generated outlier of 15 meters. Clearly, the proposed FDE mechanism is able to detect and eliminate the outlier while preserving the rest of the dynamics. As expected, the overall position accuracy is improved when using an adaptive measurement model, although a slight shift in the mean position can be seen. The latter effect can be explained by a slight mismatch of the constant and adaptive noise models in terms of an overall impact with respect to assumed noise dynamics as well as a particular satellite geometry. The provided plots are centered with respect to the RTK mean position which corresponds to the coordinates (0,0) in Fig. 2.

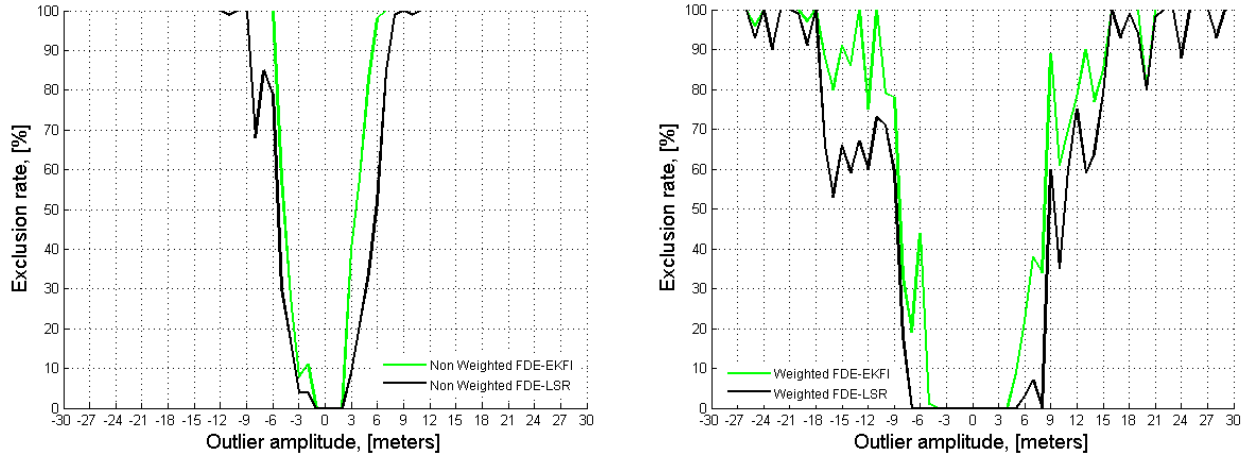
The simulation of the FDE performance has been implemented using a Monte-Carlo technique, where the true measurements were contaminated with the outliers of known amplitude. The range of the applied amplitudes was chosen from -30 till 30 meters with a step of 1 meter. For each amplitude value a random visible satellite was chosen and a random start time within the data time span was selected for the outlier injection. The outlier duration was taken to be 30 samples (15 seconds) and for each amplitude value 10 random outlier injections have been performed.

The FDE exclusion rate statistics is shown in Fig. 3 both for weighted and non-weighted LS and KF methods. In all the non-weighted approaches the $\sigma = 2$ meters was used which is reasonably close to the average residual value of 2.3 meters extracted from reference data employed for adaptive model calculation. Clearly, for reasonably large outlier amplitudes all the methods converge to the detection rate of 100%, although the performance for moderate outlier amplitudes is significantly different. Although the shown results confirm a superiority of the KF-based FDE scheme over a snapshot LS approach, the provided results represent an averaged behavior and the impact of the particular satellite is not clearly visible. Fig. 4(a) shows the fault exclusion rate for the particularly 'bad' satellite SV 3 with elevation angle during the scenario changing from 15.5 to 14 degrees and the CNo values oscillating between 30 dB-Hz and 44 dB-Hz (see Fig. 5(a)). On the other hand a performance of the fault exclusion mechanism for a 'good' satellite can be found in Fig.4(b).

The results in terms of the horizontal position error (HPE) are shown correspondingly for the LS and KF methods in Fig. 6. The results have been generated by converting the solver solution (X, Y, Z) coordinates to ENU frame and taking the RTK mean position as a zero reference. Although the FDE results in the outlier elimination in both LS and EKF, the mean HPE is shifted to larger values when the weighted measurement noise model is applied. This, however, could be explained by the a particular geometry during the scenario and inability of the weighted model to capture the range offsets along with the variance. Due to a short duration of the test scenario we were not able to sample different satellite geometries and to assess whether shift in the HPE is a systematic result or is merely a coincidence. As we have seen in Fig.2, the adaptive noise approach significantly improves the spread of the solution around the mean, although the effect of a slight shift of the mean position still needs some further investigation.

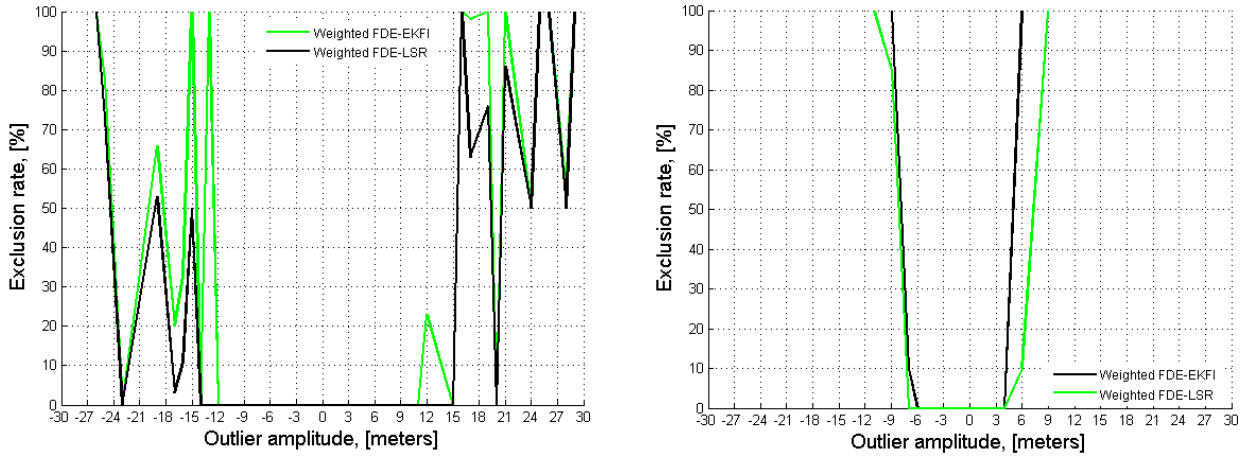
5 DISCUSSION

The results from above confirm that the FDE KF-based techniques constantly outperform the non-KF techniques when equivalent measurement noise model is used. This, however, should come at no surprise as the KF has an explicit dynam-



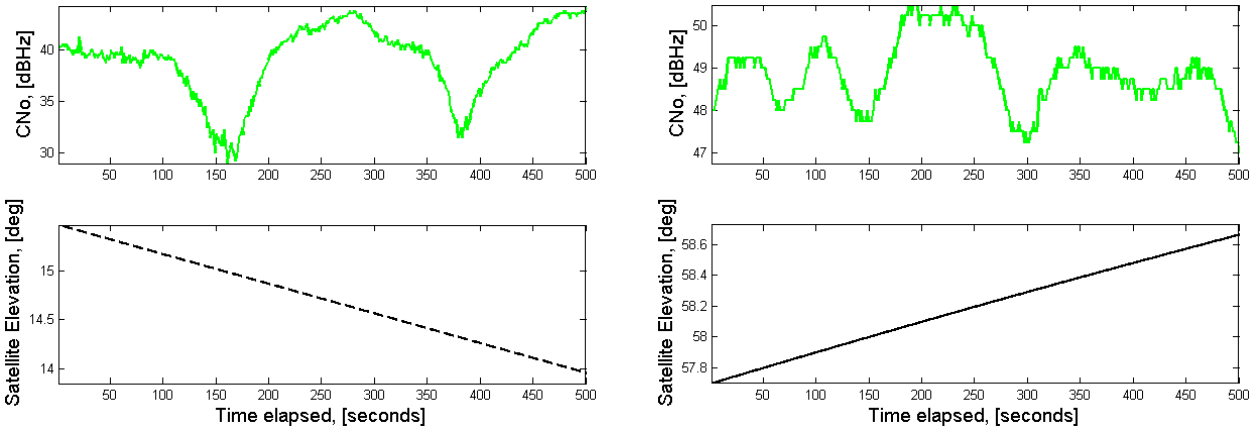
(a) (b)

Figure 3: Fault exclusion rate in Monte-Carlo simulation. The non-weighted LS and KF approach (a) and the weighted LS and KF methods (b). In both cases the elevation mask of 8° was applied and Pfa = 0.001 was used.



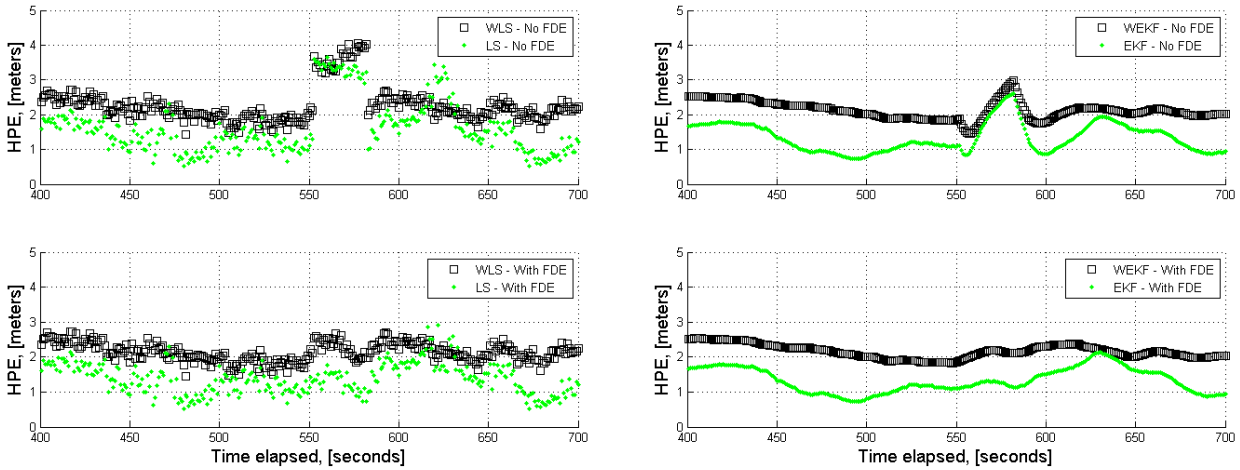
(a) (b)

Figure 4: Fault exclusion rate in Monte-Carlo simulation of the weighted LS and KF methods for a 'bad' satellite SV3 (a) and a 'good' satellite SV8 (b). In both cases the elevation mask of 8° was applied and Pfa = 0.001 was used. The results of the fault exclusion and detection for non-weighted methods are not shown as they are similar to those in Fig.3(a).



(a) (b)

Figure 5: Carrier-to-Noise Ratio Density and the satellite elevation for a 'bad' satellite SV3 (a) and a 'good' satellite SV8 (b).



(a) (b)

Figure 6: HPE for the non-weighted and weighted LS (a) and non-weighted and weighted KF (b). In both cases the elevation mask of 8° was applied and Pfa = 0.001 was used with 15 meter outlier imposed on SV8.

ics model (in this case a static position model is assumed) which fits nicely to the scenario and the results could be worse when the KF process model does not match the true dynamics. This is, fortunately, is not a problem for the system in mind when the inertial sensors are employed within the prediction step. In this case the process dynamics is based not on the assumptions on expected motion models, but rather on a true dynamics provided by a direct strapdown integration. The inertial unit would provide a true short-term stable dynamics and the FDE mechanism would benefit from this information.

Interestingly that for a 'bad' satellite SV3 (Fig. 5(a)) the outliers have been found even when they have not been added manually (not shown). Although we have tried to select a scenario with few inherent outliers, a poor intrinsic performance of the SV3 caused the outlier detector to be triggered. What is even more interesting, is that for the weighted noise model the fault exclusion rate is extremely noisy and converges to 100% only for the outlier amplitudes close to 30 meters. This, however, is a direct result of the adaptive noise model which assigns increased measurement noise variance for SV3. This results in the situations that even the outliers of a significant magnitude can be still considered 'within' the measurement statistics of the satellite with bad CNo and therefore are not excluded. On the other hand, the performance of the FDE for a relatively good satellite is similar for both weighted and unweighted (not shown) approaches where a difference in the width of the non-detection region can be, at least till some extend, explained by the average measurement noise mismatch. Obviously, most of the noise visible in Fig. 3(b) is effectively caused by the 'bad' satellite as, when it is selected for the outlier injection, the outlier exclusion can easily fail due to extremely conservative statistics assigned to this particular measurement. Here comes an important conclusion that a direct adoption of the weighted measurement model, although results in accuracy improvement, could also lead to the failure of the FDE or similar mechanisms. The reason is, that while the FDE scheme tries to check the measurement consistency using given statistics, the weighted approach tries to adjust the statistics to fit the observations, and, therefore, serves exactly an opposite purpose.

Obviously, a constant amplitude outlier injection can be hardly considered as the most representative sensor failure approach. For example, a performance of the KF-based techniques can become much worse for the ramp-like scenarios, where small amplitude and prolonged duration offsets in one of the measurements, when initially undetected, could force the filter to drift significantly from the true estimate. On the other hand, the step-like outliers form a fairly representative error model when one considers multipath effects in maritime environments [9].

6 CONCLUSIONS AND FUTURE WORK

Within this preliminary work we have successfully demonstrated an application of the FDE mechanism for outlier elimination both in snapshot and KF-based algorithms. The work is implemented within a framework of the integrated PNT unit with an additional integrity monitoring functionality. The FDE mechanism provides consistent improvements in terms of the horizontal accuracy both in LS and RBE methods. An interesting behavior of the proposed FDE mechanism has been

noticed for the methods with CNo-dependent measurement noise models, where the fault detection rate had shown worse performance compared to that of the non-adaptive methods. The performance of the proposed methods was evaluated using real measurements from the vessel combined with Monte-Carlo simulation for the outlier detection. Further work is planned in extending the presented concepts for the Doppler shift measurements, verification of the approach with IMU/GNSS augmented systems and improvement of the associated FDE algorithms for complex sensor fusion scenarios. A special attention has to be paid to a challenging problem of the memory effects in RBE schemes and associated upper error bounds.

7 BIBLIOGRAPHY

[1] International convention for the safety of life at sea (SOLAS) convention. chapter V, safety of navigation., 2002.

[2] Development of an e-navigation strategy, 2007.

[3] Kai Borre and Gilbert Strang. *Algorithms for Global Positioning*. Wellesley-Cambridge Press, 2010.

[4] M. S. Grewal and A. P. Andrews. *Kalman Filtering. Theory and Practice using MATLAB, 2nd edition*. John Wiley & Sons, Inc. New York, 2001.

[5] Heidi Kuusniemi. *User-Level Reliability and Quality Monitoring in Satellite-Based Personal Navigation*. PhD thesis, Tampere University of Technology, Tampere, September 2005.

[6] David Minkwitz, Stephan Schlüter, and Jamila Beckheinrich. Integrity assessment of a maritime carrier phase based GNSS augmentation system. In *ION GNSS 2010*, Portland, USA, September 2010. Institute of Navigation.

[7] Bradford W. Parkinson and Penina Axelrad. Autonomous GPS integrity monitoring using the pseudorange residual. *Navigation: Journal of The Institute of Navigation*, 35(2):255–274, 1988.

[8] Mark G. Petovello. *Real-Time Integration of a Tactical-Grade IMU and GPS for High-Accuracy Positioning and Navigation*. PhD thesis, University of Calgary, Calgary, Alberta, April 2002.

[9] Samuel Joseph Ryan. *Augmentation of DGPS for Marine Navigation*. PhD thesis, Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, Canada, October 2002.

[10] P.J.G. Teunissen. *GPS for Geodesy*, chapter Quality control and GPS. Springer, 1998.

[11] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, September 2005.

[12] Todd Walter, Todd Walter, and Per Enge. Weighted RAIM for precision approach, 1995.

[13] Jian-Guo Wang. Test statistics in Kalman filtering. *Journal of Global Positioning Systems*, 7(1):81–90, 2008.

[14] A. Wieser and F. Brunner. An extended weight model for GPS phase observations. *Earth Planet Space*, 52:777–782, 2000.